

CONVERGENCE IN DISTRIBUTION OF SOME SELF-INTERACTING DIFFUSIONS: THE SIMULATED ANNEALING METHOD

SÉBASTIEN CHAMBEU AND ALINE KURTZMANN

ABSTRACT. We study some self-interacting diffusions living on \mathbb{R}^d solutions to:

$$dX_t = dB_t - g(t)\nabla V(X_t - \bar{\mu}_t)dt$$

where $\bar{\mu}_t$ is the empirical mean of the process X , V is an asymptotically strictly convex potential and g is a given function, not increasing too fast to the infinity or constant. The authors have already proved that the ergodic behavior of X is strongly related to g . We go further and, using the simulated annealing method, we give some conditions for the convergence in distribution of X toward X_∞ (which law is related to the global minima of V). We also investigate the case $g(t) = 1$.

1. INTRODUCTION

In [3], the authors have obtained some conditions for both the pointwise ergodicity and the almost sure convergence of some self-interacting diffusions. We will go further in the study of such processes. The aim of this paper is to obtain some conditions first, for the convergence in probability, and second, the convergence in distribution of the self-interacting diffusion X defined by

$$(1.1) \quad dX_t = dB_t - g(t)\nabla V(X_t - \bar{\mu}_t)dt, \quad X_0 = x$$

where B is a standard Brownian motion and $\bar{\mu}_t$ denotes the empirical mean of X :

$$(1.2) \quad \bar{\mu}_t = \frac{1}{r+t} \left(r\bar{\mu} + \int_0^t X_s ds \right), \quad \bar{\mu}_0 = \bar{\mu}.$$

Here μ is an initial probability measure on \mathbb{R}^d , $\bar{\mu}$ denotes the mean of μ and $r > 0$ is an initial weight.

This paper deals with the well-known theory of simulated annealing, which has been developed since the 80's. For physical systems, an important question is to find the globally minimum energy states of the system. Experimentally, the ground states are reached by a procedure, called the chemical annealing. Let us explain the procedure. One first melts a substance and then cools it slowly enough to pass through the freezing temperature. If the temperature decreases too fast, then the system does not end up into a ground state, but in a local (but not global) minimum. On the other hand, if the temperature decreases too slowly, then the system approaches the ground states very

A.K. is partially supported by the Swiss National Science Foundation grant 200020-112316/1.

slowly. The competition between these two effects determines the optimal speed of cooling, that is the annealing schedule.

The study of the simulated annealing involves the theory of (homogeneous and) non-homogeneous Markov chains and diffusion processes, large deviation theory, spectral analysis of operator and singular perturbation theory. Pioneer work has been done by Freidlin and Wentzell [6]. The initial problem consists in finding the global minima of a function U . Indeed, one has to study the Markov process X^ε in \mathbb{R}^d given by the Langevin-type Markov diffusion (we emphasize that $\varepsilon = \varepsilon(t)$)

$$(1.3) \quad dX_t^\varepsilon = \varepsilon dB_t - \nabla U(X^\varepsilon) dt.$$

X^ε may be considered as a perturbation of the trajectory X^0 of the dynamical system $\frac{dX_t^0}{dt} = -\nabla U(X_t^0)$. Let us explain briefly the model. If the temperature ε is almost constant for a sufficiently large amount of time, then the process X^ε and the fixed temperature process behave approximatively the same at the end of that time-interval. Denote by \min the set of all the global minima of U . The optimal annealing schedule (that is, ε), for the convergence criterion $\mathbb{P}_x(X_t^\varepsilon \in \min) \rightarrow 1$ as t goes to $+\infty$, was first determined by Hajek [7] for a finite state space. Later, Chiang, Hwang and Sheu [4] studied the convergence rate of the latter probability via the large deviations of the transition density of X^ε . This rate is actually strongly related to the spectral gap of the invariant measure of X^ε .

Note, that Chiang, Wang and Sheu were one of the firsts to show the convergence of the algorithm of the simulated annealing, in the case $\varepsilon(t)^2 = k/\log t$ for k large enough. Later, Royer [16] obtained the same result for $k > \Lambda$, where Λ is related to the second eigenvalue λ_2^ε of the corresponding infinitesimal generator. Moreover, Hwang and Sheu [12] established (by probabilistic methods) the existence of $\Lambda := \lim_{\varepsilon \rightarrow 0} -\varepsilon^2 \log \lambda_2^\varepsilon$. Finally, Holley and Stroock [10] initiated an other method and proved, in the discrete case, the convergence of the simulated annealing algorithm via Sobolev's inequality. They went further in their study with Kusuoka [8]. After that, Miclo [15] proved, by using some functional inequalities, that the free energy (that is the relative entropy of the distribution of the process at time t with respect to the invariant probability at that time t) satisfies a differential inequality, implying (under some decreasing evolution of the temperature to zero) the convergence of the process to the global minima of the potential.

A natural question arises: what happens if the temperature, that is ε , decreases too fast to zero? Then, the potential can freeze in a local minimum (the “choice” of this minimum depends on the initial condition) and therefore the process converges to this local minimum. We won't consider this case here.

First, we study the process $Y_t := X_t - \bar{\mu}_t$, which satisfies the following SDE

$$(1.4) \quad \begin{cases} dY_t = dB_t - g(t)\nabla V(Y_t)dt - Y_t \frac{dt}{r+t}; & Y_0 = x - \bar{\mu}; \\ d\bar{\mu}_t = Y_t \frac{dt}{r+t}. \end{cases}$$

We will adapt the simulated annealing method to Y . We will prove that, depending on g , either the process Y converges in distribution (and not in probability) towards a variable which is concentrated on the global minima of V or converges in probability to a random

variable, which support is \mathbb{R}^d . Suppose that $V = W + \chi$, where W is strictly convex and χ is a compactly supported function. Define $\text{osc}(\chi) := \sup_x \chi(x) - \inf_x \chi(x)$. Denote by $(m_i)_{1 \leq i \leq n}$, respectively $(M_i)_{1 \leq i \leq p}$, the local minima, respectively maxima and saddle points of V . We have proved in [3], that there exist some nonnegative constants a_i such that $\sum_i a_i = 1$ and for all continuous bounded f : $\frac{1}{t} \int_0^t f(Y_s) ds \xrightarrow{a.s.} \sum a_i f(m_i)$. Here are the main results of the present work, corresponding to two classes of functions g . First, we prove, for a logarithmic g , that X converges under a condition of symmetry on the critical points of V :

Theorem 1.1. *Suppose $\infty > \lim_{t \rightarrow \infty} g(t)^{-1} \log G(t) = k > \max\{2 \text{osc}(\chi), d/4\}$. Then the process X converges in distribution to $Y_\infty + \int_0^\infty Y_s \frac{ds}{r+s}$ if and only if V is such that $\sum_{1 \leq i \leq p} a_i m_i = 0$.*

Second, we show that, for a constant g , X converges under a condition of symmetry on V :

Theorem 1.2. *Suppose that $\lim_{t \rightarrow \infty} g(t) = 1$. Then $\bar{\mu}_t$ converges in probability and X_t converges in probability to $\bar{\mu}_\infty + Y_\infty$, where Y_∞ has the normalized distribution density $e^{-2V(x)}/Z$, if and only if $\int x e^{-2V(x)} dx = 0$. Else X diverges.*

The paper is organized in the following way. In Section 2, we recall the notations and some results of [3]. Section 3 is devoted to the study of the process Y . In particular, we prove the convergence in distribution of Y towards global minima thanks to the simulated annealing method. Afterwards, we deduce in Section 4 some conditions for the convergence of the self-interacting process X . Finally, we will study the constant case $g \equiv 1$ in Section 5.

2. NOTATION, HYPOTHESIS AND FORMER RESULTS

We denote by G the function $G(t) := \int_0^t g(s) ds$ and G^{-1} is its generalized inverse: $G^{-1}(t) := \inf\{u \geq 0; G(u) \geq t\}$. In the whole following, (\cdot, \cdot) denotes the Euclidian scalar product. We denote by $\mathcal{P}(\mathbb{R}^d)$ the set of probability measures on \mathbb{R}^d .

In the sequel, we suppose that $V : \mathbb{R}^d \rightarrow \mathbb{R}_+$ is such that:

- (1) (*regularity and positivity*) $V \in \mathcal{C}^2(\mathbb{R}^d)$ and $V \geq 0$;
- (2) (*convexity*) $V = W + \chi$ where χ is a compactly supported function and there exists $c > 0$ such that $\nabla^2 W(x) \geq cId$ and $\nabla \chi$ is Lipschitz (with the constant $\tilde{C} > 0$);
- (3) (*growth*) there exist $a, b > 0$ such that for all $x \in \mathbb{R}^d$, we have

$$(2.1) \quad \Delta V(x) \leq a + bV(x) \text{ and } \lim_{|x| \rightarrow \infty} \frac{|\nabla V(x)|^2}{V(x)} = \infty.$$

We also assume that V has a finite number of critical points. Let $Max = \{M_1, M_2, \dots, M_p\}$ be the set of saddle points and local maxima of V and $Min = \{m_1, m_2, \dots, m_n\}$ be the set of the local minima of V . We assume that $\forall i, \forall \xi \in \mathbb{R}^d$, $(\nabla^2 V(m_i)\xi, \xi) > 0$ and for all M_i , $\nabla^2 V(M_i)$ admits a negative eigenvalue.

We also assume that the application $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ belongs to $\mathcal{C}^1(\mathbb{R}_+)$ and, without any loss of generality, that $g(0) > 0$. In the following, we will consider the cases $g(t) = k \log t$ and $g(t) = 1$.

Remark 2.1. *If $\lim_{t \rightarrow \infty} g(t) = \infty$, then for all $T > 0$, we have that $G^{-1}(t+T) - G^{-1}(t) \xrightarrow[t \rightarrow \infty]{} 0$.*

We have already shown that the SDE (1.1) studied admits a unique global strong solution:

Proposition 2.2. *([3] proposition 3.3) For any $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}(\mathbb{R}^d)$ and $r > 0$, there exists a unique global strong solution $(X_t, t \geq 0)$ of (1.1).*

Theorem 2.3. *(theorem 5.6) Suppose that $\lim_{t \rightarrow \infty} g(t) = \infty$ and $\lim_{t \rightarrow \infty} g'(t)/g^2(t) = 0$. Then, a.s. the normalized occupation measure of Y converges weakly to a convex combination of Dirac measures taken in the critical points of V : there exist $a_i \geq 0$ such that $\sum_i a_i = 1$ and for all continuous bounded function f , we have a.s.*

$$\frac{1}{t} \int_0^t f(Y_s) ds \xrightarrow[t \rightarrow \infty]{} \sum_{i=1}^n a_i f(m_i).$$

3. ASYMPTOTIC BEHAVIOR OF Y

Consider the time-changed process $Z_t := Y_{G^{-1}(t)}$, satisfying the following SDE

$$(3.1) \quad dZ_t = \frac{1}{\sqrt{g \circ G^{-1}(t)}} dW_t - \left(\nabla V(Z_t) + \frac{Z_t}{(r + G^{-1}(t))g \circ G^{-1}(t)} \right) dt, \quad Z_0 = z,$$

where W is a Brownian motion such that $\frac{1}{\sqrt{g \circ G^{-1}(t)}} dW_t$ has the same law as $B_{G^{-1}(t)}$ ¹. We identify $1/\sqrt{g \circ G^{-1}(t)}$ as the temperature in the simulated annealing model. So, define $\varepsilon^2(t) := \frac{1}{g \circ G^{-1}(t)}$. The process Z reads

$$(3.2) \quad dZ_t = \varepsilon(t) dB_t - \left(\nabla V(Z_t) + \frac{Z_t \varepsilon(t)^2}{r + G^{-1}(t)} \right) dt.$$

In this part, we suppose that $\frac{\log G(t)}{g(t)}$ is bounded and that $g'(t)/g(t)^2$ converges to zero.

3.1. Tightness. We begin to prove that the law of the process Z is a tight family of measures.

Lemma 3.1. *(Chiang-Hwang-Sheu) There exist some $R, c > 0$ such that*

$$\mathbb{E} \left(V(Z_t) \mathbb{1}_{\{V(Z_t) \geq R\}} \right) \leq \frac{c}{g \circ G^{-1}(t)}.$$

¹The Wiener processes B_t and W_t are not the same, but this does not matter because we are only interested in the probability distribution.

Proof. Let us first exhibit a constant γ such that $\mathbb{E}(V(Z_t)) \leq \mathbb{E}V(Z_0) + \gamma t$. By Itô's formula, we have

$$\frac{d}{dt}\mathbb{E}V(Z_t) = -\mathbb{E}\left[\left(\nabla V(Z_t), \frac{Z_t \varepsilon(t)^2}{(r + G^{-1}(t))}\right) + |\nabla V(Z_t)|^2 - \frac{\Delta V(Z_t)}{2} \varepsilon(t)^2\right].$$

As $|\nabla V|^2 - \Delta V$ is bounded from below, the first assertion follows. Now, we adapt the proof given by Dufflo [5]. The growth condition (2.1) implies that there exists r_0 such that, for $V(x) \geq r_0$, we get $|\nabla V(x)|^2 \geq 2cV(x)$ where $c > 0$ is a constant. Let ϕ be a nonnegative and nondecreasing function of class \mathcal{C}^2 , $\phi : \mathbb{R} \rightarrow [0, 1]$ such that $\phi(x) = 0$ for $x \leq r_0$, $\phi(x) = 1$ for $x \geq R$ where $r_0 < R < \infty$. Remark, that the continuous function $\nabla(\phi \circ V) = (\phi' \circ V)\nabla V$ is bounded and consider the application $\psi := (\phi \circ V)V$. We apply Itô's formula to the function $x \mapsto \psi(x)$ and we get

$$\begin{aligned} d\psi(Z_t) &= \varepsilon(t)(\nabla\psi(Z_t), dW_t) + \frac{\varepsilon(t)^2}{2}\Delta\psi(Z_t)dt \\ &\quad - \left(\frac{\psi(Z_t)}{V(Z_t)} + V(Z_t)\phi' \circ V(Z_t)\right) \left(|\nabla V(Z_t)|^2 + \frac{\varepsilon(t)^2}{r + G^{-1}(t)}(Z_t, \nabla V(Z_t))\right) dt. \end{aligned}$$

Let $\alpha(t) := \mathbb{E}[\phi \circ V(Z_t)V(Z_t)]$. By the first assertion, α is well defined. Due to the compact support of ϕ' and because we can decompose $\Delta\psi$ as

$$\begin{aligned} \Delta\psi(z) &= \phi'' \circ V(z)V(z)|\nabla V(z)|^2 + \phi' \circ V(z)V(z)\Delta V(z) \\ &\quad + 2\phi' \circ V(z)|\nabla V(z)|^2 + \phi \circ V(z)\Delta V(z), \end{aligned}$$

there exists $\tilde{C} > 0$ such that $\Delta\psi \leq \tilde{C}((\phi \circ V)V + 1)$. So, we get the bound

$$\int_t^{t+h} \varepsilon(s)^2 \mathbb{E}(\Delta\psi(Z_s)) ds \leq \tilde{C}(G^{-1}(t+h) - G^{-1}(t)) + \tilde{C} \int_t^{t+h} \alpha(s) \varepsilon(s)^2 ds.$$

On the other hand, we have the lower bound

$$\begin{aligned} \mathbb{E} &[(\phi \circ V(Z_t) + V(Z_t)\phi' \circ V(Z_t)) |\nabla V(Z_t)|^2] \\ &\geq 2c\mathbb{E}[(\phi \circ V(Z_t) + V(Z_t)\phi' \circ V(Z_t)) V(Z_t)] \geq 2c\alpha(t). \end{aligned}$$

For $r \geq r_0$ large enough, if $V(z) \geq r$ then $(z, \nabla V(z)) \geq 0$, so we get

$$\mathbb{E}((\phi \circ V(Z_t) + V(Z_t)\phi' \circ V(Z_t)) (Z_t, \nabla V(Z_t))) \geq 0.$$

Therefore, the preceding Itô's formula leads to

$$\begin{aligned} \alpha(t+h) - \alpha(t) &\leq -2c \int_t^{t+h} \alpha(s) ds + \frac{\tilde{C}}{2} \int_t^{t+h} \alpha(s) \varepsilon(s)^2 ds \\ &\quad + \frac{\tilde{C}}{2}(G^{-1}(t+h) - G^{-1}(t)), \end{aligned}$$

Letting h go to zero, this yields to $\alpha'(t) \leq -2c\alpha(t) + \frac{\tilde{C}}{2}(\alpha(t) + 1)\varepsilon(t)^2$. Choose r large enough, so that we have

$$\alpha'(t) \leq -c\alpha(t) + \frac{\tilde{C}}{2}\varepsilon(t)^2.$$

In order to solve this inequation, let $\alpha(t) := \beta(t)e^{-ct}$. We have $\beta(t) \leq \int_0^t \frac{\tilde{C}e^{cs}}{2}\varepsilon(s)^2 ds$. As $\lim_{t \rightarrow +\infty} \frac{g'(t)}{g^2(t)} = 0$, this yields to $\alpha(t) \leq C_1 \varepsilon(t)^2/c$ where C_1 is a positive constant independent of c . To conclude, we just need to remark that $\mathbb{1}_{\{V \geq R\}} \leq \phi \circ V$. \square

Corollary 3.2. *The time marginal laws of the processes Y and Z are tight.*

Proof. The previous result implies

$$\mathbb{E}V(Z_t) = \mathbb{E}(V(Z_t)\mathbb{1}_{V(Z_t) < R}) + \mathbb{E}(V(Z_t)\mathbb{1}_{V(Z_t) \geq R}) \leq R + \frac{c}{g(0)}.$$

So, for a constant $A > 0$, there exists a compact set K such that $\{V \leq A\} \subset K$ and

$$\mathbb{P}(Z_t \in K) \geq \mathbb{P}(V(Z_t) \leq A) \geq 1 - \frac{\mathbb{E}(V(Z_t))}{A} \xrightarrow{A \rightarrow \infty} 1. \quad \square$$

3.2. Convergence in distribution towards the global minima of V . Remember that $\varepsilon^2(t) := \frac{1}{g \circ G^{-1}(t)}$ and let $a(t) := (r + G^{-1}(t))\varepsilon^{-2}(t)$. The process Z reads

$$(3.3) \quad dZ_t = \varepsilon(t)dB_t - \nabla V_t(Z_t)dt$$

where we have defined $V_t(x) := V(x) + \frac{|x|^2}{a(t)}$. Actually, we will prove that this non-homogeneous Markov process converges in distribution to its “invariant” probability measure. Of course, if we suppose that $a(t) \equiv a$ and $\varepsilon(t) \equiv \varepsilon$, then the convergence in distribution is well-known.

Let $L_{t,\varepsilon}$ be the operator defined by $L_{t,\varepsilon} := \frac{1}{2}\varepsilon^2\Delta - (\nabla V_t, \nabla)$ and $P^{t,\varepsilon}$ the associated semigroup. Define

$$(3.4) \quad \Pi_{t,\varepsilon}(dx) := \frac{1}{Z(t,\varepsilon)}e^{-2\varepsilon^{-2}V_t(x)}dx,$$

where $Z(t,\varepsilon) := \int_{\mathbb{R}^d} e^{-2\varepsilon^{-2}V_t(x)}dx$. As $|\nabla V_t|^2 - \Delta V_t$ is bounded from below, the theory of Schrödinger operator implies that $L_{t,\varepsilon}$ is self-adjoint in $L^2(\Pi_{t,\varepsilon})$ and admits a spectral gap. Furthermore, when $|\nabla V_t|^2 - \Delta V_t$ goes to the infinity as $|x| \rightarrow \infty$ then the spectrum of $L_{t,\varepsilon}$ is discrete: $0 = \lambda_1(t,\varepsilon) < -\lambda_2(t,\varepsilon) < \dots$. Heuristically, when the time is of order $e^{\varepsilon^{-2}\lambda_2}$, the transition density has a nice lower bound and the process is very close to the “invariant probability” $\Pi_{t,\varepsilon}$. So, our main goal is to compute the convergence of the latter probability measure when t goes to the infinity. What is more, as the subspace corresponding to the first eigenvalue $\lambda_1(t,\varepsilon)$ is composed by the constant functions, we find

$$\lambda_2(t,\varepsilon) = \inf \left\{ \int |\nabla \phi|^2 d\Pi_{t,\varepsilon}; \quad \text{Var}_{\Pi_{t,\varepsilon}}(\phi) = 1, \phi \in \mathcal{C}^\infty(\mathbb{R}^d) \right\}.$$

So, our first aim is to compute the eigenvalue λ_2 and study its behavior when $t \rightarrow \infty$ (that is $\varepsilon \rightarrow 0$).

Consider for a while $\Pi_{\infty,\varepsilon} := \frac{1}{Z(\varepsilon)}e^{-2\varepsilon^{-2}V(x)}dx$, where $Z(\varepsilon) := \int_{\mathbb{R}^d} e^{-2\varepsilon^{-2}V(x)}dx$. Let $\min = \{m_1, \dots, m_q\}$ be the set of the global minima of V . Hwang [11] has established that $\Pi_{\infty,\varepsilon}$ converges weakly when ε converges to zero and described the limit:

Lemma 3.3. ([11]) *When t goes to the infinity, the probability measure $\Pi_{\infty, \varepsilon(t)}$ converges weakly to*

$$(3.5) \quad \Pi_0 := \frac{1}{\sum_{1 \leq i \leq q} (\det \nabla^2 V(m_i))^{-1/2}} \sum_{1 \leq i \leq q} (\det \nabla^2 V(m_i))^{-1/2} \delta_{m_i}.$$

Remark 3.4. *We use here a weaker form than the result of Hwang, who has actually proved the convergence of $\Pi_{\infty, \varepsilon(t)}$ for a more general set \min .*

We can now state and prove the

Proposition 3.5. *Suppose that (asymptotically) $g(t) > \frac{d}{4} \log t$. The probability measure $\Pi_{t, \varepsilon(t)}$ converges weakly to Π_0 as t goes to the infinity, where Π_0 is defined by (3.5).*

Proof. Recall, that $\varepsilon^2(t)a(t) = 2(r + G^{-1}(t))$ and

$$Z(t, \varepsilon(t)) = \int_{\mathbb{R}^d} e^{-2\varepsilon^{-2}(t)V(x)} e^{-2\frac{|x|^2}{\varepsilon^2(t)a(t)}} dx.$$

Let K be the compact set $K := \{x | V(x) \leq 1\}$. There exists a constant $A > 0$ such that $K \subset B(0, A)$. Then, on one hand, we get the upper bound

$$\int_{K^c} e^{-2\varepsilon^{-2}(t)V(x)} e^{-2\frac{|x|^2}{\varepsilon^2(t)a(t)}} dx \leq \int_{\mathbb{R}^d} e^{-2\varepsilon^{-2}(t)V(x)} e^{-2\frac{|x|^2}{\varepsilon^2(t)a(t)}} dx = \frac{(2\pi a(t)\varepsilon^2(t))^{d/2}}{e^{2\varepsilon^{-2}(t)}}.$$

On the other hand, we obtain similarly the lower bound

$$\int_K e^{-2\varepsilon^{-2}(t)V(x)} dx \geq \int_K e^{-2\varepsilon^{-2}(t)V(x)} e^{-2\frac{A}{a(t)\varepsilon^2(t)}} dx.$$

By Laplace's method, we have (see [11, 17])

$$\int_K e^{-2\varepsilon^{-2}(t)V(x)} dx \underset{t \rightarrow +\infty}{\sim} \sum_{i=1}^q (\pi \varepsilon^2(t))^{d/2} (\det \nabla^2 V(m_i))^{-1/2},$$

where $(m_i)_i$ are the global minima of V (they form a finite set). On the other hand, as $g(t) > \frac{d}{4} \log t$, we have that $G^{-1}(t)$ goes to the infinity and so

$$a(t)^{d/2} e^{-2\varepsilon^{-2}(t)} = (G^{-1}(t)g \circ G^{-1}(t))^{d/2} e^{-2g \circ G^{-1}(t)} \xrightarrow[t \rightarrow \infty]{} 0.$$

As a consequence, we find the asymptotic equivalence

$$Z(t, \varepsilon(t)) \underset{t \rightarrow +\infty}{\sim} \sum_i (\pi \varepsilon^2(t))^{d/2} (\det \nabla^2 V(m_i))^{-1/2}.$$

By the same method, if ϕ is a continuous function, with compact support containing only the global minimum m_1 , we find

$$\int_{\mathbb{R}^d} \phi(x) e^{-2\varepsilon^{-2}(t)V(x)} e^{-2\frac{|x|^2}{a(t)\varepsilon^2(t)}} dx \underset{t \rightarrow +\infty}{\sim} (\pi \varepsilon^2(t))^{d/2} (\det \nabla^2 V(m_1))^{-1/2} \phi(m_1).$$

This concludes the proof and also give the explicit form of Π_0 . □

To prove that Z converges in distribution towards the global minima of V , we follow the approach initiated by Holley and Stroock [10], Holley, Kusuoka and Stroock [8] and Miclo [15], using some functional inequalities. We suppose in the following that $g \circ G^{-1}(t)$ is asymptotically equivalent to $k \log(1+t)$ for k large enough. The remainder of the Section is the following. First, we will show that the measures $(\Pi_{t,\varepsilon(t)}, t \geq 0)$ satisfy a logarithmic Sobolev inequality. This will prove useful for the convergence of the free energy to zero. After that, we show that Z converges in distribution to a random variable of law Π_0 .

Definition 3.6. *The measure μ satisfies the logarithmic Sobolev inequality, with the constant C , denoted $LSI(C)$, if for all function $h \in L^2(\mu)$, we have*

$$\int h^2 \log h^2 d\mu - \left(\int h^2 d\mu \right) \log \left(\int h^2 d\mu \right) \leq C \int |\nabla h|^2 d\mu.$$

Let $p(s, x, t, y)$ denote the density of the semi-group corresponding to the non-homogeneous Markov process Z . We remind that we supposed $V = W + \chi$ and c is the convexity constant of W .

Lemma 3.7. *The family of probability measures $(\Pi_{t,\varepsilon(t)}, t \geq 0)$ satisfies a logarithmic Sobolev inequality $LSI(C(t))$, where $C(t) = 2e^{\frac{2}{\varepsilon(t)^2} \text{osc}(\chi)} / c$ (and $\text{osc}(\chi) = \sup \chi - \inf \chi$).*

Proof. The point is to use the celebrated Bakry-Emery Γ_2 -criterion (see [1]). Indeed, to the operator $L_{t,\varepsilon(t)}$, we can associate the operator “carré du champ”: for all function $f, g \in \mathcal{C}^\infty$

$$(3.6) \quad \Gamma_t^V(f, g) := \frac{1}{2} (L_{t,\varepsilon(t)}(fg) - fL_{t,\varepsilon(t)}g - gL_{t,\varepsilon(t)}f).$$

Then, we define the operator Γ_2^V as

$$(3.7) \quad \Gamma_2^V(t)(f) := \frac{1}{2} (L_{t,\varepsilon(t)}\Gamma_t^V(f, f) - 2\Gamma_t^V(f, L_{t,\varepsilon(t)}f)).$$

The Γ_2 -criterion asserts that if there exists a positive constant C such that $\Gamma_2^V \geq C\Gamma_t^V$, then $\Pi_{t,\varepsilon(t)}$ satisfies a logarithmic Sobolev inequality, $LSI(2/C)$. An easy calculation, for all function f of class \mathcal{C}^∞ , leads to

$$\Gamma_t^V(f, f) = \frac{\varepsilon(t)^2}{2} |\nabla f|^2$$

and

$$\Gamma_2^V(t)(f) = \frac{\varepsilon(t)^2}{2} (\nabla f, \nabla^2 V \nabla f) + \frac{\varepsilon(t)^4}{4} \|\nabla^2 f\|^2 + \frac{\varepsilon(t)^2}{2a(t)} |\nabla f|^2.$$

Recall, the decomposition $V = W + \chi$ where W is strictly convex with a constant c and χ is a compactly supported function. We apply Bakry and Emery’s criterion to the function $W + |x|^2/a(t)$ and we get that $\Gamma_2^W(t)(f) \geq c\Gamma_t^W(f)$. So, the probability measure $e^{-2\varepsilon^{-2}(t)(W(x)+|x|^2/a(t))}/Z$ satisfies the inequality $LSI(2/c)$. We conclude, by Holley and Stroock’s perturbation lemma [9], that the measure $\Pi_{t,\varepsilon(t)}$ satisfies a $LSI(C(t))$ inequality with $C(t) \leq 2e^{\frac{2}{\varepsilon(t)^2} \text{osc} \chi} / c$. \square

Let us denote by $(P_s^{t,\varepsilon}, s \geq 0)$ the \mathcal{C}^0 semigroup corresponding to the non-homogeneous Markov process Z . For any function f and any probability measure μ (such that f is μ -integrable), we note $\langle f \rangle_\mu := \int_{\mathbb{R}^d} f d\mu$.

Lemma 3.8. *The probability measure $\Pi_{t,\varepsilon}$ admits a spectral gap: there exists a constant $\lambda_2(\varepsilon) > 0$ such that for all $s \geq 0$, all continuous $f \in L^2(\Pi_{t,\varepsilon})$*

$$\|P_s^{t,\varepsilon} f - \Pi_{t,\varepsilon} f\|_{L^2(\Pi_{t,\varepsilon})} \leq e^{-\lambda_2(t,\varepsilon)s} \text{Var}_{\Pi_{t,\varepsilon}}(f).$$

Proof. As $\Pi_{t,\varepsilon(t)}$ satisfies the inequality $LSI(C(t))$, we get the spectral inequality with constant $\lambda_2(t, \varepsilon(t)) > 0$. \square

We want to use the previous functional inequalities in order to prove the convergence of Z_t (and thus Y_t) towards the global minima of V .

Definition 3.9. *The free energy (up to an additive constant), or relative Kullback information, of a measure P absolutely continuous with respect to Π is: $H(P|\Pi) := \int dP \log \frac{dP}{d\Pi}$. Equivalently, if we suppose that P (respectively Π) has a density p (respectively π) with respect to the Lebesgue measure λ , then we define*

$$(3.8) \quad H(p|\Pi) := \int p \log \frac{p}{\pi} d\lambda.$$

Proposition 3.10. *For all initial t_0, x_0 , we get*

$$\begin{aligned} \frac{d}{dt} H(p(t_0, x_0, t, \cdot) | \Pi_{t,\varepsilon(t)}) &\leq -\frac{2}{C(t)} \varepsilon(t)^2 H(p(t_0, x_0, t, \cdot) | \Pi_{t,\varepsilon(t)}) \\ &\quad - 4\dot{\varepsilon}(t) \varepsilon(t)^{-3} \int p(t_0, x_0, t, \cdot) (V_t - \langle V_t \rangle_{\Pi_{t,\varepsilon(t)}}) d\lambda \\ &\quad + \frac{2}{\varepsilon(t)^2} \int p(t_0, x_0, t, \cdot) (\dot{V}_t - \langle \dot{V}_t \rangle_{\Pi_{t,\varepsilon(t)}}) d\lambda. \end{aligned}$$

Proof. We adapt the proof of Holley and Stroock [10] (and Miclo [15]). In order to shorten notation, let $p_t := p(t_0, x_0, t, \cdot)$ be the distribution law of the process Z_t , knowing that $Z_{t_0} = x_0$ (the existence of a density p_t follows from Girsanov theorem). Recall, that the family of probability measures $(\Pi_{t,\varepsilon(t)}, t \geq 0)$ satisfies a family of Sobolev logarithmic inequalities ($LSI(C(t)), t \geq 0$) and $\Pi_{t,\varepsilon(t)}(dx) = \pi_{t,\varepsilon(t)}(x) \lambda(dx)$. Let us introduce

$$h_t := \sqrt{\frac{p_t}{\pi_{t,\varepsilon(t)}}},$$

satisfying in particular $\int h_t^2 d\Pi_{t,\varepsilon(t)} = 1$. So, by Lemma 3.7, there exists a constant $C(t)$ such that

$$(3.9) \quad H(p_t | \Pi_{t,\varepsilon(t)}) = \int p_t \log \frac{p_t}{\pi_{t,\varepsilon(t)}} d\lambda \leq C(t) \int |\nabla h_t|^2 d\Pi_{t,\varepsilon(t)}.$$

Computing the gradient function of h_t , we get

$$\nabla h_t = \frac{\sqrt{p_t}}{2\pi_{t,\varepsilon(t)}} \left(\frac{\nabla p_t}{p_t} + 2 \frac{\nabla V_t}{\varepsilon(t)^2} \right).$$

We put this last estimate in the preceding inequality to get

$$(3.10) \quad H(p_t | \Pi_{t,\varepsilon(t)}) \leq \frac{C(t)}{4} \int p_t \left| \frac{\nabla p_t}{p_t} + 2 \frac{\nabla V_t}{\varepsilon(t)^2} \right|^2 d\lambda.$$

Moreover, the time-derivative of the free energy H is

$$(3.11) \quad \frac{d}{dt} H(p_t | \Pi_{t,\varepsilon(t)}) = \int \dot{p}_t \log \frac{p_t}{\pi_{t,\varepsilon(t)}} d\lambda - \int p_t \frac{\dot{\pi}_{t,\varepsilon(t)}}{\pi_{t,\varepsilon(t)}} d\lambda.$$

Our strategy is to find an upper bound for both terms of (3.11). Kolmogorov forward equation reads

$$(3.12) \quad \dot{p}_t = \frac{1}{2} \varepsilon(t)^2 \Delta p_t + (\nabla p_t, \nabla V_t) + \operatorname{div}(V_t) p_t = \nabla \cdot \left(\frac{1}{2} \varepsilon(t)^2 \nabla p_t + p_t \nabla V_t \right).$$

We have the following estimates on $\pi_{t,\varepsilon(t)}$:

$$(3.13) \quad \frac{\dot{\pi}_{t,\varepsilon(t)}}{\pi_{t,\varepsilon(t)}} = 4 \frac{\dot{\varepsilon}(t)}{\varepsilon(t)^3} (V_t - \langle V_t \rangle_{\pi_{t,\varepsilon(t)}}) - \frac{2}{\varepsilon(t)^2} (\dot{V}_t - \langle \dot{V}_t \rangle_{\pi_{t,\varepsilon(t)}}),$$

$$(3.14) \quad \frac{\nabla \pi_{t,\varepsilon(t)}}{\pi_{t,\varepsilon(t)}} = -2 \frac{\nabla V_t}{\varepsilon(t)^2}.$$

Putting all the pieces together, integrating by parts and using (3.10), we get

$$\begin{aligned} \int \dot{p}_t \log \frac{p_t}{\pi_{t,\varepsilon(t)}} d\lambda &= \int \log \frac{p_t}{\pi_{t,\varepsilon(t)}} \nabla \cdot \left(\frac{1}{2} \varepsilon(t)^2 \nabla p_t + p_t \nabla V_t \right) d\lambda \\ &= - \int \left(\frac{\nabla p_t}{p_t} - \frac{\nabla \pi_{t,\varepsilon(t)}}{\pi_{t,\varepsilon(t)}}, \frac{1}{2} \varepsilon(t)^2 \nabla p_t + p_t \nabla V_t \right) d\lambda \\ &= - \frac{\varepsilon(t)^2}{2} \int p_t \left| \frac{\nabla p_t}{p_t} + 2 \frac{\nabla V_t}{\varepsilon(t)^2} \right|^2 d\lambda \\ &\leq - \frac{2}{C(t)} \varepsilon(t)^2 H(p_t | \Pi_{t,\varepsilon(t)}). \end{aligned}$$

On the other hand, we obtain for the second term of (3.11):

$$\int p_t \frac{\dot{\pi}_{t,\varepsilon(t)}}{\pi_{t,\varepsilon(t)}} d\lambda = 4 \frac{\dot{\varepsilon}(t)}{\varepsilon(t)^3} \int p_t (V_t - \langle V_t \rangle_{\pi_{t,\varepsilon(t)}}) d\lambda - \frac{2}{\varepsilon(t)^2} \int p_t (\dot{V}_t - \langle \dot{V}_t \rangle_{\pi_{t,\varepsilon(t)}}) d\lambda.$$

Putting all the pieces together in (3.11) leads to the result. \square

Lemma 3.11. *For all $t \geq 0$, the quantity $\langle |x|^2 \rangle_{\Pi_{t,\varepsilon(t)}}$ is bounded.*

Proof. Let K be the compact set $K := \{x | V(x) \leq \eta\}$ where η is a given positive constant. As $\Pi_{t,\varepsilon(t)}$ converges weakly to Π_0 we only need to prove that $\langle |x|^2 \mathbb{1}_{K^c} \rangle_{\Pi_{t,\varepsilon(t)}}$ is bounded.

For any $1 \geq \varepsilon(t)^2$, we have

$$\begin{aligned} \int_{K^c} |x|^2 e^{-2\varepsilon^{-2}(t)V_t(x)} dx &\leq \int_{K^c} |x|^2 e^{-2V(x)} e^{-2V(x)(\varepsilon^{-2}(t)-1)} dx \\ &\leq e^{-2\eta(\varepsilon^{-2}(t)-1)} \int_{K^c} |x|^2 e^{-2V(x)} dx \end{aligned}$$

But, as proved in Proposition 3.5, we have the asymptotic equivalence

$$Z(t, \varepsilon(t)) \underset{t \rightarrow \infty}{\sim} \sum_i (\pi \varepsilon^2(t))^{d/2} (\det \nabla^2 V(m_i))^{-1/2},$$

which implies $\langle |x|^2 \mathbb{1}_{K^c} \rangle_{\Pi_{t, \varepsilon(t)}} \leq \tilde{C} \varepsilon^{-d}(t) e^{-2\eta \varepsilon^{-2}(t)} \xrightarrow{t \rightarrow \infty} 0$. \square

In order to conclude the convergence of the free energy to zero, we will use an easy result of Miclo [15]

Lemma 3.12. (*Miclo, lemma 6*) *Let $f : [0, \infty[\rightarrow \mathbb{R}_+$ be a continuous function such that a.s.*

$$f'(t) \leq \alpha_t - \beta_t f(t),$$

where α and β are two continuous non-negative functions such that $\int^\infty \beta_t dt = \infty$ and $\lim_{t \rightarrow \infty} \alpha_t / \beta_t = 0$. Then $\lim_{t \rightarrow \infty} f(t) = 0$.

Proof. One has to adapt Gronwall's lemma. Let $g(t) = f(t) \exp\left(\int_0^t \beta_s ds\right)$, which is continuous. We get a.s.

$$g'(t) \leq \alpha_t \exp\left(\int_0^t \beta_s ds\right).$$

We therefore obtain that

$$g(t) \leq g(0) + \int_0^t \alpha_s \exp\left(\int_0^s \beta_u du\right) ds.$$

Consequently, for all $t_0 \geq 0$, we have

$$f(t) \leq f(0) e^{-\int_0^t \beta_u du} + e^{-\int_0^t \beta_u du} \left\{ \int_0^{t_0} \alpha_s e^{\int_0^s \beta_u du} ds + \int_{t_0}^t \alpha_s e^{\int_0^s \beta_u du} ds \right\}.$$

Let $\eta > 0$. We choose t_0 such that for all $t \geq t_0$, we have $\alpha_s \leq \eta \beta_s$. We thus find an upper bound, for any $t \geq t_0$, for the last term of the preceding inequality:

$$e^{-\int_0^t \beta_u du} \int_{t_0}^t \eta \beta_s e^{\int_0^s \beta_u du} ds \leq e^{-\int_0^t \beta_u du} \eta \left(e^{\int_0^t \beta_u du} - e^{\int_0^{t_0} \beta_u du} \right) \leq \eta.$$

As the two first terms go to zero when t goes to the infinity, we get $\limsup_{t \rightarrow \infty} f(t) \leq \eta$. \square

Corollary 3.13. *If $\lim_{t \rightarrow +\infty} \frac{\alpha'_t}{\alpha_t \beta_t} - \frac{\beta'_t}{\beta_t^2} = 0$ and $\lim_{t \rightarrow +\infty} \frac{\alpha_t}{\beta_t} e^{\int_0^t \beta_s ds} = +\infty$, then we have asymptotically (when $t \rightarrow \infty$)*

$$\int_0^t \alpha_s e^{\int_0^s \beta_u du} ds \sim \frac{\alpha_t}{\beta_t} e^{\int_0^t \beta_s ds}.$$

Proof. Integrating by parts, we have

$$\int_0^t \alpha_s e^{\int_0^s \beta_u du} ds = \frac{\alpha_t}{\beta_t} e^{\int_0^t \beta_s ds} - \frac{\alpha_0}{\beta_0} - \int_0^t \left(\frac{\alpha'_s}{\alpha_s \beta_s} - \frac{\beta'_s}{\beta_s^2} \right) \alpha_s e^{\int_0^s \beta_u du} ds. \quad \square$$

Theorem 3.14. *Suppose that $\varepsilon^2(t) = k/\log(t)$, with $\infty > k > 2 \operatorname{osc}(\chi)$. Then, for all initial t_0, x_0 , the free energy $H(p(t_0, x_0, t, \cdot) | \Pi_{t, \varepsilon(t)})$ converges to 0 (as $t \rightarrow \infty$).*

Proof. Let $t_0 \geq 0$ and $x_0 \in \mathbb{R}^d$. Consider the process Z_t , solution to the SDE

$$dZ_t = \varepsilon(t) dW_t - \left(\nabla V(Z_t) + \frac{Z_t}{a(t)} \right) dt, \quad Z_{t_0} = x_0.$$

The result of Proposition 3.10 can be rewritten in the following way:

$$\begin{aligned} \frac{d}{dt} H(p_t | \Pi_{t, \varepsilon(t)}) &\leq -\frac{2}{C(t)} \varepsilon(t)^2 H(p_t | \Pi_{t, \varepsilon(t)}) + \frac{2}{\varepsilon(t)^2} (\mathbb{E} \dot{V}_t(Z_t) - \langle \dot{V}_t \rangle_{\Pi_{t, \varepsilon(t)}}) \\ &\quad - 4\dot{\varepsilon}(t) \varepsilon(t)^{-3} \left(\mathbb{E} V_t(Z_t) - \langle V_t \rangle_{\Pi_{t, \varepsilon(t)}} \right). \end{aligned}$$

As $V(x) \geq c|x|^2$ out of a compact set and Z is tight (by Corollary 3.2), we have $\mathbb{E} V_t(Z_t) = O(1)$. Because $t \mapsto a(t)$ is nondecreasing while $t \mapsto \varepsilon(t)$ is nonincreasing, and as $\dot{V}_t(x) = -\frac{\dot{a}(t)}{a(t)^2} |x|^2$, the two terms $\mathbb{E}(\dot{V}_t(Z_t))$ and $\langle \dot{V}_t \rangle_{\Pi_{t, \varepsilon(t)}}$ are nonpositive. So, it only remains to find an upper bound for $\langle \dot{V}_t \rangle_{\Pi_{t, \varepsilon(t)}}$. Indeed, by Lemma 3.11, there exist $M_1, M_2 > 0$ such that

$$\frac{d}{dt} H(p_t | \Pi_{t, \varepsilon(t)}) \leq -\frac{2}{C(t)} \varepsilon(t)^2 H(p_t | \Pi_{t, \varepsilon(t)}) + M_1 \frac{\dot{\varepsilon}(t)}{\varepsilon(t)^3} + M_2 \frac{\dot{a}(t)}{\varepsilon(t)^2 a(t)^2}.$$

We easily compute the time-derivative of $a(t)$:

$$\begin{aligned} \frac{\dot{a}(t)}{a(t)^2 \varepsilon^4(t)} &= -\frac{2\dot{\varepsilon}(t)}{\varepsilon^3(t)(r + G^{-1}(t))} + \frac{1}{(g \circ G^{-1}(t))(r + G^{-1}(t))^2 \varepsilon^2(t)} \\ &= \frac{1}{kt(r + G^{-1}(t))} + \frac{\log t}{k(g \circ G^{-1}(t))(r + G^{-1}(t))^2}. \end{aligned}$$

As $G^{-1}(t)$ is a nondecreasing function and because of the hypothesis on k , the term $\frac{C(t)}{kt(r + G^{-1}(t))}$ converges to 0 when t goes to the infinity. For the second term, we recall that $\log G(t)/g(t)$ is bounded by assumption: there exist two positive constants m, M such that $mg(t) \leq \log G(t) \leq Mg(t)$. So, we get $mg(t) \leq \log(tg(t)) = \log t + \log g(t)$, what naturally implies that $g(t) = O(\log t)$ and then $G(t) \leq tg(t) = o(t^2)$. So,

$$G(t)^{2 \operatorname{osc}(\chi)/k} \log G(t) / (kg(t)(r + t)^2) \xrightarrow[t \rightarrow \infty]{} 0.$$

Lemma 3.12 asserts that if ε satisfies

$$\int_1^\infty \varepsilon(t)^2 \frac{dt}{C(t)} = \infty \text{ and } \frac{\dot{\varepsilon}(t)}{\varepsilon^5(t)} \xrightarrow{t \rightarrow \infty} 0,$$

then $\lim_{t \rightarrow \infty} H(p_t | \Pi_{t, \varepsilon(t)}) = 0$. We meet the required conditions with $\varepsilon^2(t) = k / \log t$. \square

Remark 3.15. *The constant k is not optimal here, because we have used the perturbation lemma of Holley and Stroock for the estimation of the logarithmic Sobolev constant.*

Lemma 3.16. *The speed of convergence of $H(p(t_0, x_0, t, \cdot) | \Pi_{t, \varepsilon(t)})$ toward 0 is $\frac{(\log t)^3}{t^{2-2 \operatorname{osc}(\chi)/k}}$.*

Proof. As we satisfy the hypothesis of Corollary 3.13, the speed of convergence is $\frac{\alpha_t}{\beta_t}$, with $\alpha_t = (r + G^{-1}(t))^{-2} + (t(r + G^{-1}(t)))^{-1}$ and $\beta_t = t^{-2 \operatorname{osc}(\chi)/k} / \log t$. So, we get

$$\frac{\alpha_t}{\beta_t} = \frac{t^{2 \operatorname{osc}(\chi)/k} \log t}{(r + G^{-1}(t))^2} + \frac{\log t}{t^{1-2 \operatorname{osc}(\chi)/k} (r + G^{-1}(t))}.$$

As, up to a multiplicative constant, $g \circ G^{-1}(t) = \log t$, we get that $g(t) = O(\log t)$ and $G^{-1}(t)$ is of the order of $t / \log t$. So α_t / β_t is asymptotically of the order of $(\log t)^3 t^{2 \operatorname{osc}(\chi)/k-2}$. \square

Remark 3.17. *It is known since the work of Freidlin and Wentzell [6] (see [5], Chap.5), that the Gibbs measure $\Pi_{t, \varepsilon(t)}$ satisfies a large deviation principle. Therefore, the speed of convergence of $\Pi_{t, \varepsilon(t)}$ toward Π_0 is $e^{-\log t / 2k}$.*

Corollary 3.18. *Suppose that $\varepsilon^2(t) = k / \log(t)$, with $\infty > k > \max\{2 \operatorname{osc}(\chi), d/4\}$. Then Z and Y converge in distribution to a random variable, which law is Π_0 (3.5).*

Proof. The Kullback information $H(p_t | \Pi_{t, \varepsilon(t)})$ estimates the distance between p_t and $\Pi_{t, \varepsilon(t)}$ in the following way: $\|p_t - \Pi_{t, \varepsilon(t)}\|_{TV}^2 \leq 2H(p_t | \Pi_{t, \varepsilon(t)})$, where $\|\cdot\|_{TV}$ denotes the total variation norm (see [10]). The result follows because $\Pi_{t, \varepsilon(t)}$ converges weakly to Π_0 and the total-variation norm characterizes the weak convergence of measures. \square

Remark 3.19. *We emphasize that if $\lim_{t \rightarrow \infty} g(t)^{-1} \log G(t) = k$, with k not large enough, then the previous result is false and Y does not converge toward the global minima of V (except if each local minimum of V is a global minimum). Indeed, in the simulated annealing theory, if ε decreases too fast to zero, then the process Y freezes in a local minimum, the choice of the minimum depending on the initial value $Y_0 = x - \bar{\mu}$.*

4. THE PROCESS X : CONVERGENCE IN DISTRIBUTION

We give here sufficient conditions for the convergence of X . As usual, we begin to work with the process $Y_t = X_t - \bar{\mu}_t$. In order to link this section with the preceding one, we recall that $\varepsilon(t)^2 = (g \circ G^{-1}(t))^{-1} = k / \log t$. So, we consider only functions g such that (asymptotically) $\log G(t) = kg(t)$. By Theorem 2.3, there exist $a_i \geq 0$ such that $\sum a_i = 1$ and for all continuous bounded f : $\frac{1}{t} \int_0^t f(Y_s) ds \rightarrow \sum a_i f(m_i)$.

Theorem 4.1. *Suppose that $\infty > \lim_{t \rightarrow \infty} g(t)^{-1} \log G(t) = k > \max\{2 \operatorname{osc}(\chi), d/4\}$. Then one of the following holds:*

- (1) *If V is a function such that $\sum_{1 \leq i \leq n} a_i m_i = 0$, then X_t converges in distribution to $Y_\infty + \int_0^\infty Y_s \frac{ds}{r+s}$;*
- (2) *Else, X_t diverges.*

Proof. We wonder whether $\int_0^t Y_s \frac{ds}{r+s}$ converges in distribution or not. Remark, that

$$X_t = Y_t + \int_0^t \frac{ds}{r+s} Y_s = Y_t + \bar{\mu}_t.$$

Lemma 3.1 ensures that the duality for the weak convergence is true for functions bounded by V , so in particular for $f(x) = x$. Suppose that V is such that $\sum_{1 \leq i \leq n} a_i m_i = 0$. Then,

we know that $\frac{1}{t} \int_0^t Y_s ds \xrightarrow{a.s.} 0$ and it remains to find the rate of convergence in order to conclude the proof. But, Benaïm and Schreiber [2] (Theorem 1) have proved that, for an asymptotic pseudotrajectory (in probability) Y , the speed of convergence of the mean of the normalized occupation measure of Y is the same as the speed of the pseudotrajectory. This means that the speed of convergence of the normalized occupation measure of the time-changed process $Y_{G^{-1}(t)}$ is $G^{-1}(1+t) - G^{-1}(t)$. Integrating by parts, we obtain

$$\frac{1}{t} \int_0^t Y_s ds = \frac{1}{tg(t)} \int_0^{G(t)} Y_{G^{-1}(u)} du + \frac{1}{t} \int_0^{G(t)} du \frac{g' \circ G^{-1}(u)}{(g \circ G^{-1}(u))^3} \int_0^u Y_{G^{-1}(s)} ds.$$

The first right-hand term converges to 0 because $G(t) \leq tg(t)$. It remains to prove the convergence of the second term. We have that (because, up to a multiplicative constant, $g \circ G^{-1}(u) = \log(2+u)$)

$$\begin{aligned} & \frac{1}{t} \int_0^{G(t)} du \frac{g' \circ G^{-1}(u)}{(g \circ G^{-1}(u))^3} s [G^{-1}(s+T) - G^{-1}(s)] ds \\ & \leq \frac{G(t)(t - G^{-1}(G(t) + T))}{tg(t)} + \frac{1}{t} \int_0^t \frac{G^{-1}(s+T) - G^{-1}(s)}{g \circ G^{-1}(s)} ds < \infty. \end{aligned}$$

So, if V is a function such that $\sum_{1 \leq i \leq n} a_i m_i \neq 0$, then $\int_0^t Y_s ds$ does not converge. Suppose that $\sum_{1 \leq i \leq n} a_i m_i = 0$. Let $U_t := \bar{\mu}_t$ and $V_t := X_t$. We conclude, because the celebrated Slutsky theorem asserts that for two sequences of \mathbb{R}^d -valued random variables (U_t) and (V_t) , with $U_t \xrightarrow[t \rightarrow \infty]{(d)} U$ and $|U_t - V_t| \xrightarrow[t \rightarrow \infty]{\mathbb{P}} 0$, then $V_t \xrightarrow[t \rightarrow \infty]{(d)} U$. \square

Remark 4.2. *The condition on g means that, asymptotically, $g(t) = k \log t$ for k large enough.*

5. THE CASE $g(t) = 1$

We give the proof for \mathbb{R} but it is easily reproduced in \mathbb{R}^d . We suppose now that $g(t) = 1$. In order to study the behavior of the process X solution of

$$dX_t = dB_t - V'(X_t - \bar{\mu}_t)dt,$$

we introduce the process $Y_t := X_t - \bar{\mu}_t$, solution of the SDE

$$dY_t = dB_t - V'(Y_t)dt - \frac{Y_t}{r+t}dt.$$

We also introduce the Kolmogorov process Z solution to

$$dZ_t = dB_t - V'(Z_t)dt.$$

This process is a positive recurrent diffusion. Denote by γ its invariant probability measure, $\gamma(dx) = \frac{e^{-2V(x)}}{\int_{\mathbb{R}} e^{-2V(y)}dy}dx$, and $\bar{\gamma}$ is the mean of γ . For all $h \in L^1(\gamma)$ we have, with an exponential speed of convergence,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t h(Z_s)ds = \int_{\mathbb{R}} h d\gamma \quad a.s.$$

Lemma 5.1. *The process Y converges in probability to a random variable Y_∞ of density γ when t goes to $+\infty$.*

Proof. Using that $-y(y-z) \leq \frac{z^2}{2}$ for all $y, z \in \mathbb{R}$, we get

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \mathbb{E}(Y_t - Z_t)^2 &= -\mathbb{E}(V'(Y_t) - V'(Z_t), Y_t - Z_t) - \mathbb{E}\left(\frac{Y_t}{r+t}, Y_t - Z_t\right) \\ &\leq -(c + \tilde{C})\mathbb{E}(Y_t - Z_t)^2 + \frac{1}{r+t}\mathbb{E}(Z_t^2). \end{aligned}$$

Applying Itô's formula, it is easy to prove the existence of $M > 0$ such that for all $t \geq 0$, $\mathbb{E}(Z_t^2) \leq M$. So, $\mathbb{E}(Y_t - Z_t)^2$ goes to zero. We choose the random variable Z_0 , which distribution function is γ , so that the law of Z_t is γ for all t . So, Y converges in L^2 to a random variable of law γ . \square

Lemma 5.2. *For all $A > 0$, we have $\lim_{t \rightarrow \infty} \sup_{0 \leq s \leq A} \left| \int_0^s (Y_{e^{t+u}} - \bar{\gamma}) du \right| = 0$ a.s.*

Proof. Fix $T > 0$. We note Y^T the process defined by $Y_s^T := Y_{s+T}$. It is solution to

$$dY_s^T = dB_s^T - V'(Y_s^T)ds - \frac{Y_s^T}{s+T}ds.$$

Recall, that $\bar{\gamma} = \int_{\mathbb{R}} x\gamma(x)dx$. Let $A > 0$, we have for all $s \leq A$ and $T = e^t$:

$$\int_0^s (Y_{e^{t+u}} - \bar{\gamma}) du = \int_0^{T(e^s-1)} \frac{Y_{v+T} - Z_v^T}{v+T} dv + \int_0^{T(e^s-1)} \frac{Z_v^T - \bar{\gamma}}{v+T} dv =: I + J,$$

where the process Z^T is the solution to the SDE

$$dZ_s^T = dB_s^T - V'(Z_s^T)ds, \quad Z_0^T = Y_T.$$

1) Study of J. Integrating by part, we have with $S = e^s - 1$,

$$J = \frac{TS}{TS+T} \left(\frac{1}{TS} \int_0^{TS} Z_u^T du - \bar{\gamma} \right) + \int_0^{TS} \frac{v}{(v+T)^2} \left(\frac{1}{v} \int_0^v Z_u^T du - \bar{\gamma} \right) dv.$$

The process Z satisfies the limit-quotient theorem, so the first right-hand term converges a.s. to 0 when T goes to the infinity. For the second right-hand term we have

$$\left| \int_0^{TS} \frac{v}{(v+T)^2} \left(\frac{1}{v} \int_0^v Z_u^T du - \bar{\gamma} \right) dv \right| \leq \frac{1}{T^2} \int_0^{TS} \left| \frac{1}{v} \int_0^v Z_u^T du - \bar{\gamma} \right| dv.$$

By Cesàro, J converges to 0 as T goes to the infinity.

2) Study of I. We use the estimate of the distance between the processes Y_v^T and Z_v^T for T large enough, as in Lemma 5.1. So, letting $C := 2(c - \tilde{C}) > 0$, we find the following upper bound

$$(Y_v^T - Z_v^T)^2 \leq e^{-Cv} \int_0^v \frac{e^{Cu} (Z_u^T)^2}{u+T} du.$$

Let $\sigma_\infty := \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t (Z_u^T)^2 du$. Remind, that there exists $a > 0$ such that the speed of convergence is less than e^{-at} , so that we have

$$\int_0^v \frac{e^{Cu} (Z_u^T)^2}{u+T} du = \int_0^v \frac{e^{Cu} \sigma_\infty}{u+T} du + \int_0^v \frac{e^{Cu} ((Z_u^T)^2 - \sigma_\infty)}{u+T} du =: K + L.$$

With probability 1, we obtain the upper bounds $|K| \leq \frac{e^{Cv} \sigma_\infty}{aT}$ and $|L| \leq \frac{e^{Cv}}{aT}$, implying $(Y_s^T - Z_s^T)^2 \leq \frac{\sigma_\infty + 1}{aT}$ a.s. So, $|I| \leq \frac{1}{\sqrt{T}}$ and I converges to zero as T goes to the infinity. \square

We have now to study the asymptotic behavior of $\bar{\mu}_t$. Integrating by parts, we get

$$(5.1) \quad \bar{\mu}_t = \bar{\mu}_0 + \int_0^t \frac{Y_s}{r+s} ds = \bar{\mu}_0 + \frac{t}{r+t} m_t + \int_0^t \frac{s}{(r+s)^2} m_s ds,$$

where $m_t := \frac{1}{t} \int_0^t Y_s ds$.

The proof of the main result of this section is based on the following:

Proposition 5.3. *The process m_t converges in probability to $\bar{\gamma}$. Moreover the speed of convergence of m_t toward $\bar{\gamma}$ is less than $\frac{1}{t}$.*

Proof. As $dm_t = \frac{1}{t}(Y_t - m_t)dt$, letting $n_t := m_{e^t}$, we get $\frac{d}{dt}n_t = Y_{e^t} - n_t$. Consequently, we find

$$n_{t+s} - n_t = \int_t^{t+s} (Y_{e^u} - n_u) du = \int_0^s (-n_{t+u} + \bar{\gamma}) du + \int_0^s (Y_{e^{t+u}} - \bar{\gamma}) du.$$

Let $\varepsilon_t(s) := \int_0^s (Y_{e^{t+u}} - \bar{\gamma}) du$. By Lemma 5.2, for all $A > 0$ we have $\lim_{t \rightarrow \infty} \sup_{0 \leq s \leq A} |\varepsilon_t(s)| = 0$ a.s., so n_t is an asymptotic pseudotrajectory (a.s.) for the flow generated by

$$\frac{d\psi_t(x)}{dt} = \bar{\gamma} - \psi_t(x), \quad \psi_0(x) = x.$$

As this flow admits only one limit point which is exponentially attracted, the ω -limit set of n_t is reduced to $\{\bar{\gamma}\}$. So, $m_t = n_{\log t}$ converges a.s. to $m_\infty := \bar{\gamma}$. \square

Corollary 5.4. (1) If $\bar{\gamma} = 0$, then $\bar{\mu}_t$ converges in probability as t goes to infinity.
 (2) If $\bar{\gamma} \neq 0$, then $\bar{\mu}_t$ diverges and $\lim_{t \rightarrow \infty} \frac{\bar{\mu}_t}{\log t} = \bar{\gamma}$ (in probability).

Proof. By (5.1), we get

$$\bar{\mu}_t = \bar{\mu}_0 + \frac{t}{r+t}(m_t - \bar{\gamma}) + \int_0^t \frac{s}{(r+s)^2}(m_s - \bar{\gamma})ds + \bar{\gamma}(\log(1+t/r) - r). \quad \square$$

Summarizing, we have now proved the following.

Theorem 5.5. *One of the following holds:*

- (1) If $\bar{\gamma} = 0$, then $\bar{\mu}_t$ converges in probability to $\bar{\mu}_\infty$ and X_t converges in probability to $Y_\infty + \bar{\mu}_\infty$, where the law of Y_∞ has the density γ ;
- (2) Else, X_t diverges.

6. APPENDIX

In Section 3, the optimal annealing schedule is not obtained. Actually, k should be directly related to m , as Holley, Kusuoka and Stroock [8] proved it for $a(t) \equiv \infty$. Let $z_0 \in \mathbb{R}^d$ such that $V_t(z_0) = 0$. Let K be the compact support of χ .

Definition 6.1. *The maximal height of the function V_t is the non-negative function $m(t)$ defined by*

$$(6.1) \quad m(t) := \sup\{H_t(x, z_0) - V_t(x); \quad x \in K\},$$

where $H_t(x, z) := \inf\{E_t(\gamma); \gamma \in \mathcal{C}^1([0, 1], K), \gamma(0) = x, \gamma(1) = z\}$ and $E_t(\gamma) := \sup\{V_t(\gamma(u)); u \in [0, 1]\}$.

Remark, that $m(t)$ does not depend on z_0 and so, we choose $z_0 = 0$. The function $m(t)$ corresponds to the maximum of all the minimal energies one needs to go from each point of \mathbb{R}^d to z_0 . It is positive if and only if there exists several local minima.

Lemma 6.2. *We have that $\lim_{t \rightarrow \infty} m(t) = m$, where m is the maximal height function corresponding to V .*

Proof. Let $M := \sup\{|x|^2; x \in K\}$. For all path γ , we have clearly $E_t(\gamma) \leq E(\gamma) + \frac{M}{a(t)}$. Then, by definition we get $|H_t(x, 0) - H_\infty(x, 0)| \leq \frac{M}{a(t)}$. Consequently, there exists $C > 0$ such that $|m(t) - m(\infty)| \leq \frac{C}{a(t)}$. \square

Jacquot relates in [13] the height function to the second eigenvalue of the infinitesimal generator of Y^ε (that is the constant involved in the spectral gap inequality). He proves that $\lim_{\varepsilon \rightarrow 0} \varepsilon^2 \log \lambda_2(\infty, \varepsilon) = -2m(\infty)$. So, the “critical” value of k should be m instead of $\text{osc } \chi$.

REFERENCES

- [1] BAKRY D. & EMERY M. [1985], Diffusions hypercontractives, *Sém. Prob. XIX*, Springer LNM 1123, 177-206.
- [2] BENAÏM M. & SCHREIBER S.J. [2000], Weak asymptotic pseudotrajectories for semiflows: ergodic properties, *J. Dynamics and Diff. Eq.* **12**(3).
- [3] CHAMBEU S. & KURTZMANN A. [2007], Some particular self-interacting diffusions: ergodic behavior and almost sure convergence, preprint.
- [4] CHIANG T.S., HWANG C.R. & SHEU S.J. [1987], Diffusion for global optimization in \mathbb{R}^n , *SIAM J. Control Optim.*, **25**, 737-753.
- [5] DUFLO M. [1996], *Algorithmes stochastiques*, Springer “Mathématiques & Applications 23”.
- [6] FREIDLIN M.I. & WENTZELL A.D. [1984], *Random perturbation of dynamical systems*, Springer.
- [7] HAJEK B. [1988], Cooling schedules for optimal annealing, *Math. Oper. Res.*, **13**, 311-329.
- [8] HOLLEY R.A., KUSUOKA S. & STROOCK D.W. [1989], Asymptotics of the spectral gap with applications to the theory of simulated annealing, *J. Funct. Anal.*, **83**, 333-347.
- [9] HOLLEY R. & STROOCK D. [1987], Logarithmic Sobolev Inequalities and Stochastic Ising Models, *J. Stat. Phys.* **46**, 1159-1194.
- [10] HOLLEY R.A. & STROOCK D.W. [1988], Simulated annealing via Sobolev inequalities, *Comm. Math. Phys.*, 553-569.
- [11] HWANG C.R. [1980], Laplace’s method revisited: weak convergence of probability measures, *Ann. Proba.*, **8**, 1177-1182.
- [12] HWANG C.R. & SHEU S.J. [1990], Large-time behavior of perturbed diffusion Markov processes with applications to the second eigenvalue problem for Fokker-Planck operators and simulated annealing, *Acta App. Math.*, **19**, 253-295.
- [13] JACOT S. [1992], Comportement asymptotique de la seconde valeur propre des processus de Kolmogorov, *J. Mult. An.*, **40**, 335-347.
- [14] LÉPINGLE D. [1978], Sur le comportement asymptotique des martingales locales, *Sém. Prob. (Strasbourg)* **12**, 148-161.
- [15] MICLO L. [1992], Recuit simulé sur \mathbb{R}^n . Etude de l’évolution de l’énergie libre, *Ann. Inst. H. Poincaré*, **28** (2), 235-266.
- [16] ROYER G. [1989], A remark on simulated annealing of diffusion processes, *SIAM J. Control Optim.*, **27**, 1403-1408.
- [17] ZITT P-A. [2006], *Applications d’inégalités fonctionnelles à la mécanique statistique et au recuit simulé*, thesis Université Paris X Nanterre.

Sébastien Chambeau: Laboratoire de modélisation stochastique et statistique, Université Paris Sud, Bâtiment 425, F-91405 Orsay Cedex, France. Sebastien.Chambeau@math.u-psud.fr

Aline Kurtzmann: Institut Elie Cartan, Université Henri Poincaré Nancy 1, B.P.239, F-54506 Vandœuvre-lès-Nancy Cedex, France. Aline.Kurtzmann@iecn.u-nancy.fr